

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:14 ; Search time 170.72 Seconds
(w/o alignments)
17.898 Million cell updates/sec

Title: US-09-331-631A-1_COPY_29_73
Perfect score: 252
Sequence: 1 SEFDROEYEECKRQCMQET..... RCVSQCDKRFEDDIDNSKYD 45
Scoring table: BloSUM62
Gapcost 10.0 , Gapext 0.5
searched: 195891 seqs, 6790655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR66:
1: Pir1:
2: Pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match Length	DB ID	Description			
1	77.5	30.8	509	2 S08059	alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)		
2	77.5	30.8	588	1 FWCNAB	N:Alternate names: seed storage protein		
3	72	28.6	605	2 S06398	C:Species: Gossypium hirsutum (upland cotton)		
4	70.5	28.0	566	2 S22477	C:Accession: S08059 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993		
5	67.5	26.8	47	2 JC5557	R:Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.		
6	67	26.6	810	2 T44330	A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.		
7	67	26.6	1170	2 A40558	A:Reference number: S06398		
8	67	26.6	1170	2 T45840	A:Accession: S08059		
9	64	25.4	554	2 J01763	A:Status: not compared with conceptual translation		
10	63	25.0	524	2 T28432	A:Molecule type: DNA		
11	59.5	23.6	3078	2 T29699	C:Superfamily: glycinin		
12	58	23.0	242	2 T05892	RESULT 1		
13	58	23.0	1421	2 T05892	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
14	57	23.6	316	2 G71600	Best Local Similarity 41.2%; Pred. No. 0.11;		
15	57	22.6	725	1 RWYZ94	Matches 14; Conservative 9; Mismatches 10;		
16	57	22.6	726	1 J02162	Indels 1; Gaps 1;		
17	57	22.6	1284	1 WNWZAI	RESULT 2		
18	55	21.8	305	2 B55346	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
19	55	21.8	348	2 A71601	Best Local Similarity 41.2%; Pred. No. 0.11;		
20	55	21.8	623	2 T06674	Matches 14; Conservative 9; Mismatches 10;		
21	55	21.8	795	1 HHCH08	Indels 1; Gaps 1;		
22	55	21.8	822	2 T02824	RESULT 3		
23	54.5	21.6	244	2 S44822	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
24	54.5	21.6	1021	2 T15765	Best Local Similarity 41.2%; Pred. No. 0.11;		
25	54.5	21.6	1259	2 T32901	Matches 14; Conservative 9; Mismatches 10;		
26	54.5	21.6	2715	2 T13049	Indels 1; Gaps 1;		
27	54	21.4	291	2 S62730	RESULT 4		
28	54	21.4	314	2 T02964	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
29	54	21.4	643	2 T19225	Best Local Similarity 41.2%; Pred. No. 0.11;		

ALIGMENTS

RESULT	Match	Score	DB
1	508059	77.5	3
	alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)		
	N:Alternate names: seed storage protein		
	C:Species: Gossypium hirsutum (upland cotton)		
	C:Accession: S08059 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993		
	R:Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.		
	A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.		
	A:Reference number: S06398		
	A:Accession: S08059		
	A:Status: not compared with conceptual translation		
	A:Molecule type: DNA		
	C:Superfamily: glycinin		
	RESULT 1		
	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
	Best Local Similarity 41.2%; Pred. No. 0.11;		
	Matches 14; Conservative 9; Mismatches 10;		
	Indels 1; Gaps 1;		
	RESULT 2		
	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
	Best Local Similarity 41.2%; Pred. No. 0.11;		
	Matches 14; Conservative 9; Mismatches 10;		
	Indels 1; Gaps 1;		
	RESULT 3		
	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
	Best Local Similarity 41.2%; Pred. No. 0.11;		
	Matches 14; Conservative 9; Mismatches 10;		
	Indels 1; Gaps 1;		
	RESULT 4		
	Query 5 30.8%; Score 77.5; DB 2; Length 509;		
	Best Local Similarity 41.2%; Pred. No. 0.11;		
	Matches 14; Conservative 9; Mismatches 10;		
	Indels 1; Gaps 1;		

C:Comment: This is a seed storage protein.

Query Match 30.8%; Score 77.5; DB 1; Length 588;
 Best Local Similarity 40.5%; Pred. No. 0.13; Mismatches 11; Indels 1; Gaps 1;
 Matches 15; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

Qy 2 EFDRQEVEECKRQCMOLETS-GOMRCVSCDKRFEED 38
 Db 81 RDPQRYYEECQQECRQOE-ERORPQCQRCRKRFQE 116

RESULT 3
 S06398
 alpha-globulin type A precursor - upland cotton
 N;Alternate names: seed storage protein
 C;Species: Gossypium hirsutum (upland cotton)
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
 R;Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
 A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se
 A;Reference number: S06398
 A;Accession: S06398
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-1605 <CHI>
 C;Superfamily: glycinin
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-565/Product: alpha-globulin type A #status predicted <MAT>

Query Match 28.6%; Score 72; DB 2; Length 605;
 Best Local Similarity 42.5%; Pred. No. 0.57; Mismatches 10; Indels 4; Gaps 3;
 Matches 17; Conservative 9; Mismatches 10; Indels 4; Gaps 3;

Qy 1 SEFD-RQEYEECKRQCMOLETS-GOMRCVSCDKRFEEDI 39
 Db 33 SDPPQDQRYEDCRKRC-QLETRKQTEQ--DKCEDRSBTOL 69

RESULT 4
 S22477
 vicilin precursor - cacao
 C;Species: Theobroma cacao (cacao)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
 C;Accession: S22477; S22478; S18105; S22050
 R;McHenry, L.; Fritz, P.J.
 Plant Mol. Biol. 18, 1173-1176, 1992
 A;Title: Comparison of the structure and nucleotide sequence of vicilin genes of cocoa
 A;Reference number: S22477; MUID:92288309
 A;Accession: S22477
 A;Molecule type: DNA
 A;Residues: 1-1566 <MCB>
 A;Cross-references: EMBL:X62625
 A;Accession: S22478
 A;Molecule type: mRNA
 A;Residues: 1-452 <MC2>
 A;Cross-references: EMBL:X62626
 C;Genetics:
 A;Introns: 211/1; 269/3; 296/3; 391/3; 502/1
 C;Superfamily: glycinin
 C;Keywords: seed; storage protein
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-566/Product: vicilin #status predicted <MAT>

Query Match 28.0%; Score 70.5; DB 2; Length 566;
 Best Local Similarity 32.4%; Pred. No. 0.8; Mismatches 11; Indels 1; Gaps 1;
 Matches 11; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

RESULT 5
 JC5557
 arginine/glutamate-rich 6.5K polypeptide - smooth loofah
 C;Species: Luffa cylindrica (smooth loofah)
 C;Accession: JC5557
 C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 07-May-1999
 R;Kimura, M.; Park, S.S.; Sakai, R.; Yamasaki, N.; Funatsu, G.
 Biosci. Biotechnol. Biochem. 61, 984-988, 1997
 A;Title: Primary structure of 6.5k-arginine/glutamate-rich polypeptide from the seeds
 A;Reference number: JC5557; MUID:9735743
 A;Accession: JC5557
 A;Molecule type: protein
 A;Residues: 1-47 <KIM>
 A;Experimental source: seed
 F;12-33,16-29/Disulfide bonds: #status predicted
 C;Comment: This protein is a storage protein which provides nitrogen and carbon reser
 F;12-33,16-29/Disulfide bonds: #status predicted

Query Match 26.8%; Score 67.5; DB 2; Length 47;
 Best Local Similarity 41.2%; Pred. No. 0.19; Mismatches 14; Indels 1; Gaps 1;
 Matches 14; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

Qy 5 RQEYEECKRQCMOLETS-GOMRCVSCDKRFE 37
 Db 6 RQEYEA CRVRCQVAEHGVERQRQCQQVCEKRLRE 39

RESULT 6
 T44430
 protein Pv100 [Imported] - winter squash
 C;Species: Cucurbita maxima (winter squash)
 C;Accession: T44430
 C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 R;Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
 J. Biol. Chem. 274, 2563-2570, 1999
 A;Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a sin
 A;Reference number: 222767; MUID:99107919
 A;Accession: T44430
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-810 <YAM>
 A;Cross-references: EMBL:AB019195; NID:93808061; PIDN:BA34056.1; PID:93808062

Query Match 26.6%; Score 67; DB 2; Length 810;
 Best Local Similarity 34.1%; Pred. No. 2.8; Mismatches 15; Indels 8; Gaps 2;
 Matches 15; Conservative 6; Mismatches 15; Indels 8; Gaps 2;

Qy 5 RQEYEECKRQCMOLETS-GOMRCVSCDKRFE----- 40
 Db 75 RAQEYEVCRLCRCAVQAEVGVEQQRKEEQGRGEDVD 118

RESULT 7
 TSHUPL
 thrombospondin 1 precursor - human
 C;Species: Homo sapiens (man)
 C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
 C;Accession: A26155; A34274; A30140; A25612; A05172; A12927
 R;Lawler, J.; Hynes, R.O.
 J. Cell Biol. 103, 1335-1348, 1986
 A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multipl
 A;Reference number: A26155; MUID:87057617
 A;Accession: A26155
 A;Molecule type: mRNA
 A;Residues: 1-1170 <LAW>

A;Cross-references: GB:X04665; NID:q31137; PIDN:CAA8370_1; PID:931138
A;Note: parts of this sequence, including the amino end of the mature protein, were determined by sequencing of the cDNA clone.
R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequence analysis and transcriptional activity. J. Cell Biol. 108, 729-736, 1989

A;Accession number: A34274; MUID:89291870

A;Accession type: DNA

A;Residues: 1-166 <LAH>

A;Cross-references: GB:J04835

R;Hennessy, S.W.; Frazier, A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Bornstein, P.

A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 5' untranslated region. J. Cell Biol. 108, 729-736, 1989

A;Accession number: A30140

A;Reference number: A30140; MUID:89139590

A;Molecule type: mRNA

A;Cross references: EMBL:XI4787; NID:937464; PIDN:CAA32899_1; PID:937465

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by sequencing of the cDNA clone.

R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A25812; MUID:87157592

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-397 <KOB>

A;Cross-references: GB:M25631; NID:9538353; PIDN:AAA36741_1; PID:9538354

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of the cDNA clone.

A;Accession number: A05172; MUID:86287276

A;Accession: A05172

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-397 <RC>

A;Cross-references: GB:ML1426; NID:954005; PIDN:AAA61237_1; PID:9553801

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin. J. Cell Biol. 118, 693-701, 1992

A;Reference number: A42927; MUID:92348511

A;Accession: A42927

A;Molecule type: protein

A;Residues: 987-1003 <SUN>

A;Note: Cys-992 is shown to have a free sulphydryl.

A;Complex: homotrimer, disulfide linked

C;Function:

A;Description: participates in cell migration and adhesion, and in platelet aggregation

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type 1 repeat homology

C;Keywords: beta hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer

F1-18/Domain: signal sequence #status predicted <SIG>

F1-19-17/Domain: thrombospondin 1 #status predicted <MAT>

F1-17-17/Domain: von Willebrand factor type C repeat homology <VWNC>

F1-17-17/Protein: thrombospondin 1 #status predicted <SIG>

F317-375/Domain: von Willebrand factor type C repeat homology <VWNC>

F378-429/Domain: thrombospondin type 1 repeat homology <THR1>

F434-490/Domain: thrombospondin type 1 repeat homology <THR2>

F491-547/Domain: thrombospondin type 1 repeat homology <THR3>

F551-586/Domain: EGF homology <EGF1>

F65-689/Domain: EGF homology <EGF2>

F92-928/Region: cell attachment (R-G-D) motif

F171-232/disulfide bonds: #status predicted

F248-360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

F270-274/disulfide bonds: interchain #status predicted

F610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

F1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 26.6%; Score 67; DB 1; Length 1170;
Best Local Similarity 39.5%; Pred. No. 3,9;
Matches 15; Conservative 6; Mismatches 11; Indels 6; Gaps 2;

RESULT 9
#45840
Grpase activating-like protein - Arabidopsis thaliana

QY 13 RQMOLE---TSGQMRCA-VSQCDKRFEDIDWSKY 44
| | | | | : | | | | : || || | : | | | |
Db 404 RSCDSLNRRCEGSSVQRTCHIQECDRFKQDGWSHW 441

QY 13 RQMOLE---TSGQMRCA-VSQCDKRFEDIDWSKY 44
| | | | | : | | | | : || || | : | | | |
Db 404 RSCDSLNRRCEGSSVQRTCHIQECDRFKQDGWSHW 441

N;Alternate names: protein F2K15, 210
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C;Accession: T45840
 R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,
 submitted to the Protein Sequence Database, January 2000
 A;Reference number: Z23015
 A;Accession: T45840
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-154 <XIE>
 A;Cross-references: EMBL:AL132956
 A;Experimental source: cultivar Columbia; BAC clone F2K15
 C;Genetics:
 A;Map position: 3
 A;Introns: 53/2; 98/3; 115/2; 133/2
 A;Note: F2K15.210

RESULT 10
 JO1730
 62K sucrose-binding protein precursor - soybean
 ;Species: Glycine max (soybean)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 26-Aug-1999
 C;Accession: JO1730
 R;Grimes, H.D.; Overvoorde, P.J.; Ripp, K.; Franceschi, V.R.; Hitz, W.D.
 Plant Cell, 1, 1561-1574, 1992
 A;Title: A 62-kD sucrose binding protein is expressed and localized in tissues actively
 A;Reference number: JO1730; MUID:93104680
 A;Accession: JO1730
 A;Molecule type: mRNA
 A;Residues: 1-154 <GR1>
 A;Cross-references: GB:L06038; NID:91431744; PIDN:AAB03894_1; PID:9170064
 C;Keywords: glycinin
 C;Superfamily: sugar transport
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;30-524/Product: 62K sucrose-binding protein #status predicted <MAT>

RESULT 11
 T28432
 variant-specific surface protein 1 - malaria parasite (*Plasmodium falciparum*)
 N;Alternate names: erythrocyte membrane binding protein 1 (EMP1)
 C;Species: Plasmodium falciparum
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C;Accession: T28432
 R;Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.S.
 Cell, 82, 89-100, 1995
 A;Title: The large diverse gene family var encodes proteins involved in cytoadherence an
 A;Reference number: 220487; MUID:95330813
 A;Accession: T28432
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Residues: 1-3078 <SIX>
 A;Cross-references: EMBL:L40608; NID:9886374; PID:9886375; PIDN:AAA73396.1

RESULT 14

G71600

rifin PFB1005w - malaria parasite (*Plasmodium falciparum*)

C;Species: Plasmodium falciparum

C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000

R;Gardner, M.J.; Nettelein, H.; Carrucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Perteet, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Science 282, 1126-1132, 1998A;Title: Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*.

A;Reference number: A71600; MUID:99021743

A;Accession: G71600

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-316 <GAR>

A;Cross-references: GB:AE001432; GB:AE001362; NID:93845332; PIDN: AAC71988.1; PID:9384533

A;Experimental source: clone 3D7

A;Genetics:

A;Gene: PFB1005w

Db 550 KRRNVEWLS-RLRRDIKECDK-YKEDLDKAK 579

Search completed: March 1, 2001, 15:52:17

Job time: 541 sec

RESULT 15

WM2294

Best Local Similarity 34.9%; Score 57; DB 2; Length 316;

Matches 15; Conservative 4; Mismatches 10; Indels 14; Gaps 2;

OY 3 FDHQ-----EYRECKROCMOLETSQGMRRRCVSOQCDRFEEDI 39

Db 67 FDSQTSERFEEDIDE-----RMKDKKRKCEQQCDKQEI 101

RESULT 15

WM2294

A-type inclusion protein - vaccinia virus (strain WR)

C;Species: vaccinia virus

A;Note: host Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 29-Oct-1999

C;Accession: A41701; A40825; S29908

R;De Carlos, A.; Paez, E.

Virology 185, 768-778, 1991

A;Title: Isolation and characterization of mutants of vaccinia virus with a modified

A;Reference number: A41701; MUID:92074241

A;Accession: A41701

A;Molecule type: DNA

A;Residues: 1-725 <DEC>

A;Cross-references: GB:M76371; NID:9335683; PIDN:AAA48275.1; PID:9335684

R;Amegadze, B.Y.; Sisler, J.R.; Moss, B.

Virology 186, 777-782, 1992

A;Title: Frame-shift mutations within the vaccinia virus A-type inclusion protein gene.

A;Reference number: A40825; MUID:92124754

A;Accession: A40825

A;Molecule type: DNA

A;Residues: 1-586, 'KO', 589-609, 'R', 611-618; 620-682, 'S', 684-725 <AME>

A;Cross-references: GB:M61187; NID:9335782; PIDN:AAA48321.1; PID:9335784

R;Amegadze, B.Y.

submitted to the EMBL Data Library, January 1991

A;Reference number: S29907

A;Accession: S29908

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-586, 'KO', 589-609, 'R', 611-618, 620-682, 'S', 684-725 <AM2>

A;Cross-references: EMBL:X57318; NID:962249; PIDN:CAA46574.1; PID:962241

C;Superfamily: cowpox virus A-type inclusion protein

C;Keywords: inclusion protein

Query Match 22.6%; Score 57; DB 1; Length 725;

Best Local Similarity 40.6%; Pred. No. 38; Matches 13; Conservative 11; Mismatches 6; Indels 2; Gaps 2;

OY 12 KRCQMQLETSQMRCCVQSQCDRKFEEDIDWSK 43

